Journal of Computational Physics **177,** 284–312 (2002)
doi:10.1006/jcph.2002.7011, available online at http://www.idealibrary.com on IDEAL®

Optimization of the Crystal Surface Temperature Distribution in the Single-Crystal Growth Process by the Czochralski Method

Ja Hoon Jeong and In Seok Kang[1]

Department of Chemical Engineering and Division of Mechanical Engineering, Pohang University of Science and Technology, San 31, Hyoja Dong, Pohang, 790-784, South Korea
E-mail: iskang@postech.ac.kr

Received February 22, 2001; revised January 25, 2002

The optimization of the crystal surface temperature distribution is performed for single-crystal growth in the Czochralski process. In the optimization problem, we seek an optimal solution in the sense that the index of crystalline defects is minimized while the single-crystal growth rate is maximized. In the objective function, the von Mises stress is considered a driving force that induces crystalline defects. In order to solve the optimization problem with the equality constraints given by the governing partial differential equations, the variational method is used. Based on the calculus of variations and the method of Lagrange multiplier, the Euler–Lagrange equations are derived in the form of coupled partial differential equations. They are solved by using the finite-difference method and the iterative numerical scheme proposed in this work. In order to handle inequality constraints, the penalty function method is applied. The optimal distributions of the crystal surface temperature obtained in this work may provide an insight into the optimal design of thermal surroundings, such as thermal shield configurations and heater/cooler positions.

Key Words: variational methods; coupled partial differential equations; optimization of the crystal surface temperature; thermal stress; single-crystal growth; the Czochralski method.

1. INTRODUCTION

In the Czochralski (CZ) process, the quality of single crystals is determined by crystalline defects. Recently, an attempt was made to analyze the crystalline defects by using a thermal stress model and point defects model. Tsukada *et al*. [1] reported their numerical and experimental studies on crack formation in a $LiNbO_3$ single crystal by using the von Mises

[1] Corresponding author. Fax: +82 54 279 2699.

0021-9991/02 $35.00

stress criterion. By comparing their numerical simulations with experimental observations, they proposed that the crack formation might occur when the thermal stress at the crystal surface exceeds a certain value or when the region with relatively large thermal stress extends inside the crystal. On the other hand, it has been believed that the relative supersaturation of point defects is a driving force to induce microdefects in crystalline silicon. Sinno *et al.* [2] presented their point defects model to describe the appearance of the oxidation-induced stacking-fault (OSF) ring formation created during the cooling of silicon crystals in the CZ process. They reported that the predictions of the OSF ring based on the point defects model are in excellent agreement with the empirical correlation determined from their experiments.

Because crystalline defects are strongly influenced by the crystal thermal history, the optimal distribution of the crystal temperature is the most important key in single-crystal growth by the CZ method. Many efforts have been made to suggest proper operating conditions for high-quality single-crystal growth. Bornberger and Ammon [3] reported the results of their experimental investigations into the dependence of the OSF ring in CZ-grown silicon crystals on operating conditions, where they used different heat shields. Bornside *et al.* [4] applied their integrated numerical analysis model as a design tool to find the optimum processing conditions and system configurations for the dislocation-free silicon single-crystal growth by the CZ method. They used the von Mises stress as the measure of thermal stress. By comparing the stress profiles obtained by numerical computations case by case, they suggested several design modifications to the CZ system, such as the positions of the auxiliary heater or cold sink and the geometrical configurations of the heat shield.

The purpose of this work is to propose a systematic way for optimization of crystal temperature distribution in single-crystal growth by the CZ method. In this work, the optimization problem is formulated as an inverse problem, where the policy is given rather than the system itself. In the problem, we seek an optimal solution in the sense that the index of crystalline defect is minimized while the single-crystal growth rate is maximized. In the objective function, the von Mises stress is considered the driving force that induces crystalline defects. In order to solve the optimization problem with the equality constraints given by the governing equations, the variational method is used. Based on the calculus of variations and the method of Lagrange multiplier, we derive an optimality system of equations, the so-called Euler–Lagrange equations in the form of coupled partial differential equations (PDEs). In order to handle the inequality constraints, the penalty function method is applied with the unit Heaviside step function and a penalty parameter.

The optimal distributions of the crystal surface temperature obtained from this work may provide an insight and fundamental information into the optimal design of thermal surroundings, such as thermal shield configurations and heater/cooler positions. The variational methods in this work can be extended to more complicated problems relevant to the CZ process, such as the point defect problem. In the point defect problem, the complicated state equations add some extra complexity. The rigorous computations of the point defect problem needed for consideration of the relative supersaturation of self-interstitials and vacancies will be presented in a forthcoming work by the authors.

From the mathematical viewpoint, the optimization problem in this work is an optimal boundary control problem for a distributed parameter system, described by partial differential equations. The optimization and control of distributed parameter systems has attracted much attention of researchers in applied mathematics and engineering. For studies in the

area, the calculus of variations plays the most important role as a general technique for optimization in function spaces. For the theoretical results, readers may consult excellent references, such as [5, 6]. Recent advancement in computing power has also made it possible to solve optimal control problems of practical importance, such as the optimal fluid flow control.

Since variational calculus is familiar in the finite-element analysis, most of the optimal flow control problems have been analyzed by using the finite-element methods. He *et al*. [7] performed drag optimization for flow past a circular cylinder by using oscillatory cylinder rotation, where the state equations are the time-dependent Navier–Stokes equations and the quadratic objective function involves the vorticity and a regularization term. They discretized the state equations using the finite-element approximation and applied a quasi-Newton method to the discrete control problem. They reported that their results agree closely with the quasi-optimal forcing conditions determined by parametric search. In similar ways, Berggren [8] solved the vorticity minimization problem where the flow is controlled by suction and blowing on a part of the boundary. Also, Ghattas and Bark [9] solved the optimal control problem, where steady incompressible flows are controlled by suction or injection of fluid onto portions of the boundary in order to minimize the energy dissipation. To reduce the computation time, they developed reduced Hessian sequential quadratic programming methods, which are compared with other methods.

It is noteworthy that the objective functions are usually given in terms of the control variable or its derivatives. It is because the boundary velocity of the flow is used as a means to control the flow itself for the minimization of drag force, vorticity, and dissipation energy [7–9]. On the other hand, Hou and Ravindran [10] developed a systematic way to solve an optimal control problem, where an electrically conducting fluid is controlled by using electromagnetic force. They used the electric current at the boundary as the control variable to match a desired velocity field, or to minimize the vorticity in the flow domain. However, even in this case, the objective function still includes the terms of the control variable or its derivatives because the use of the control variable, namely the normal electric current on the boundary, is minimized simultaneously. Surprisingly enough, there have been only few attempts to discuss the case in which the objective function does not include the control variable or its derivatives. One of the reasons comes from the fact that in most optimal control problems, the objective function mainly consists of two performance indices. One is about the difference between the current state and the ideal or desired state, and the other is about the magnitude of the control efforts. Both of them should be simultaneously minimized; then the optimal state would be determined by a competition between the two indices.

In some situations, however, we encounter optimization problems where the control variables or their derivatives are not included in the objective functions. The problem of temperature control to minimize thermal stress in this work provides a typical example. In such problems, as is shown later, the boundary conditions of the state and adjoint variables at the control surface exhibit quite unique features and require special treatment in the numerical implementation. Especially that is the case when the problem is to be solved by an iterative numerical scheme. In the present work, we derive an auxiliary condition for the crystal surface temperature. Then we solve the system of equations of the state and adjoint variables using the finite-difference scheme with the proposed iterative numerical scheme.

2. PROBLEM STATEMENT

2.1. The Objective Function Formulation

The generation and motion of dislocation in crystals grown by the CZ method closely correlate with the thermoelastic stresses in the growing crystal. Although the process of generation and multiplication of the dislocations is far from being fully understood yet, Jordan *et al*. [11] critically reviewed and extended previous works. They explained the dislocation formation by using the concept of resolved shear stress and critical resolved shear stress, below which no dislocation generation was assumed to occur. Maroudas and Brown [12] have used the Haasen model [13] to analyze dislocation dynamics in silicon and III–V growth. According to the Haasen model, the von Mises stress, which scales the magnitude of the deviatoric portion of thermal stress, supplies the driving force for dislocation.

In our optimization problem, we seek an optimal solution in the sense that the index of crystalline defect is minimized while the single-crystal growth rate is maximized. In the objective function, the von Mises stress is considered the driving force that induces crystalline defects. The radial uniformity of the temperature distribution in the crystal phase is also considered because the produced wafer should have uniform thermal history to keep the constant mechanical and electrical properties. The single-crystal growth rate is determined by the energy balance at the crystal–melt interface. That is, the difference between the heat flux from the interface at the crystal side and the heat flux to the interface at the melt side determines the local rate of solidification. When the heat flux from the melt phase is kept constant, the crystal growth rate is proportional to the heat transfer rate toward the crystal phase minus the constant flux from the melt.

Mathematically, the optimization problem is described by

$$\text{Min. } I(u) = \alpha_u \int_\Omega F_u(u)\,\mathrm{d}\Omega + \alpha_{\mathbf{S}} \int_\Omega F_{\mathbf{S}}(\mathbf{S})\,\mathrm{d}\Omega - \alpha_G G(u), \tag{1}$$

where u denotes the dimensionless temperature in the crystal phase, $\mathbf{S}$ the dimensionless stress tensor, and α_u, $\alpha_{\mathbf{S}}$, and α_G the positive weighting parameters corresponding to the performance measures of F_u, $F_{\mathbf{S}}$, and G to represent the radial uniformity of the temperature distribution, the magnitude of the von Mises stress, and the single-crystal growth rate, respectively. For simplicity, the crystal is assumed to be a straight cylinder, as shown in Fig. 1, where the system domain is denoted by Ω. The boundaries Γ_1, Γ_2, Γ_3, and Γ_4 are defined by

$$\begin{aligned} &\Gamma_1 : z = 0, \quad 0 \le r \le 1, \qquad \Gamma_2 : r = 1, \quad 0 < z < l, \\ &\Gamma_3 : z = l, \quad 0 \le r \le 1, \qquad \Gamma_4 : r = 0, \quad 0 < z < l, \end{aligned}$$

where l means the dimensionless aspect ratio of the crystal region. Thus we have

$$F_u = \left(\frac{\partial u}{\partial r}\right)^2, \quad F_{\mathbf{S}} = (\tau_{vM})^2, \quad G = \int_{\Gamma_1} (\mathbf{n} \cdot \nabla u)\,\mathrm{d}\Gamma, \tag{2}$$

where $\mathbf{n}$ denotes the outgoing unit normal vector. The scalar value of the von Mises stress, τ_{vM}, is defined as the second invariant of the deviatoric part of the stress tensor, $\mathbf{S}$.

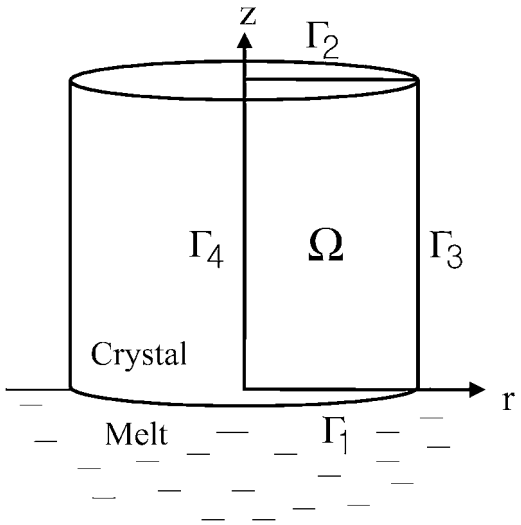

FIG. 1. Schematic diagram for the optimization of the crystal surface temperature distribution in single-crystal growth by the CZ method.

That is,

$$\tau_{vM} = \sqrt{\frac{1}{2}\mathbf{S}_d : \mathbf{S}_d}, \quad \mathbf{S}_d \equiv \mathbf{S} - \frac{1}{3}(\mathbf{S} : \mathbf{E})\mathbf{E}, \tag{3}$$

where $\mathbf{E}$ denotes the identity tensor.

It is noteworthy that the control variable u appears explicitly in the objective function when α_u has a positive value. However, when α_u is null, the objective function is given by the domain integration of another variable of $\mathbf{S}$ only. This case is discussed later from the viewpoint of numerical implementation.

2.2. Governing Equations in the Thermal Stress Problem

From the viewpoint of optimization, governing equations play the roles of equality constraints. The stress distributions are determined by Hooke's law and the force balance equations for a body in static equilibrium [14]. Their dimensionless forms are given by

$$\mathbf{D} = \frac{1}{2}((\nabla \mathbf{v}) + (\nabla \mathbf{v})^T), \tag{4}$$

$$\mathbf{S} = \frac{1}{(1+\nu)}\mathbf{D} + \frac{\nu}{(1+\nu)(1-2\nu)}(\mathbf{D} : \mathbf{E})\mathbf{E} - \frac{\beta}{(1-2\nu)}(u-1)\mathbf{E}, \tag{5}$$

$$\nabla \cdot \mathbf{S} = 0, \tag{6}$$

where $\mathbf{v}$ represents the displacement vector, $\mathbf{D}$ the strain tensor, ν the Poisson ratio, and β the dimensionless thermal expansion coefficient. When the data given in the literature [4] are used, we have $\nu = 0.25$ and $\beta = 0.01$ in the silicon single-crystal growth. We define the dimensionless temperature as $u = (T - T_{am})/(T_m - T_{am})$, where T_m and T_{am} denote the melting temperature and the ambient temperature around the crystal, respectively. It is

assumed that the magnitude of dimensionless displacement caused by the stress is much less than unity, so that their effects on the system geometry can be neglected. The gravitational force is also neglected in the calculation of the stress field. We have the traction-free conditions for all boundaries since the top and bottom plates and the side of the crystal phase are free. That is,

$$\mathbf{n} \cdot \mathbf{S} = 0 \quad \text{on } \Gamma_1, \Gamma_2, \Gamma_3. \tag{7}$$

On the other hand, the dimensional governing equation for the temperature field in the crystal is given by

$$\frac{\partial T_S}{\partial \tilde{t}} + \tilde{\mathbf{u}}_S \cdot \tilde{\nabla} T_S = \alpha_S \tilde{\nabla}^2 T_S,$$

where $\tilde{\mathbf{u}}_S$ consists of the translational velocity (pulling rate) and the rotational velocity and α_S is the thermal diffusivity. By using the fact that the convective heat transfer due to the pulling is very small in comparison with the conductive heat transfer and that the temperature profile is axisymmetric, we can obtain the simplified governing equation for the quasi-steady state in the form of Laplace equation. That is,

$$\nabla^2 u = 0, \tag{8}$$

which has the boundary conditions of

$$u = 1 \text{ on } \Gamma_1, \quad u = 0 \text{ on } \Gamma_2. \tag{9}$$

On the other hand, u is not specified along Γ_3. As is shown later, the appropriate condition of u along Γ_3 is obtained from the optimality condition.

3. CALCULUS OF VARIATIONS

The calculus of variations provides a necessary condition as an optimality system of equations. The method of Lagrange multipliers essentially gives a set of conditions necessary to find the optimal solution of equality-constrained optimization problems. This is done by converting the constrained problem to an equivalent unconstrained problem with the help of certain unspecified parameters called Lagrange multipliers. In this section, the calculus of variations and the method of Lagrange multipliers are briefly reviewed as preliminaries.

For concise but general discussion, we use the operator expressions and we adopt also the general notations rather than the notations for specific variables. In this section, u is for the state variable that is to be controlled at the boundary and s is for other state variable in the system. We consider the optimization problem given by

$$\text{Min. } I(u) = \alpha_u \int_\Omega F_u(u)\, \mathrm{d}\Omega + \alpha_s \int_\Omega F_s(s)\, \mathrm{d}\Omega, \tag{10}$$

subject to

$$L(u) = 0, \tag{11}$$

$$M(s) + N(u) = 0, \tag{12}$$

where α_u, α_s are the positive weighting parameters and L, M, N are partial differential operators. The system domain Ω is bounded by Γ_e and Γ_n. They are called the essential boundary and the natural boundary, respectively, since the state variable u is specified on Γ_e but free on Γ_n. On the other hand, the other state variable, s, is specified at every boundary and its distribution in the domain is influenced by u via (12). If we find the condition of u on Γ_n, the system becomes deterministic and can be solved. Therefore, the important thing is to find the distribution of u on Γ_n to minimize the given objective function. For the specific problem considered in this study, u corresponds to the dimensionless temperature u and s to the stress-related variable, such as the von Mises stress.

By using the method of Lagrange multipliers, we introduce the augmented objective function, given by

$$I_a = I + I_u + I_s, \tag{13}$$

where

$$I_u = \langle \hat{u}, L(u) \rangle, \quad I_s = \langle \hat{s}, M(s) + N(u) \rangle. \tag{14}$$

Here we define the inner product as $\langle u, v \rangle = \int_\Omega uv \, d\Omega$ for arbitrary scalar functions of u and v. Lagrange multipliers of $\hat{u}$ and $\hat{s}$, called the adjoint variables, have spatial distributions. As can be seen in (13), the augmented objective function, I_a, has three contributions: one corresponds to the original objective function and the others are related to the equality constraints in the optimization problem. By taking the variation to both sides of (13), we have the optimality condition, given by

$$\delta I_a = \delta I + \delta I_u + \delta I_s = 0. \tag{15}$$

First, let us consider δI. Since the system boundary is fixed, we have

$$\delta I = \alpha_u \int_\Omega \delta F_u \, d\Omega + \alpha_s \int_\Omega \delta F_s \, d\Omega. \tag{16}$$

We suppose that

$$F_u = F_u(u, \mathbf{u}_1, \mathbf{u}_2), \quad F_s = F_s(s), \tag{17}$$

where $\mathbf{u}_1$ and $\mathbf{u}_2$ are the derivatives of the unknown function u, defined by

$$\mathbf{u}_1 = \nabla u, \quad \mathbf{u}_2 = \nabla \nabla u. \tag{18}$$

By using the chain rule, we have

$$\delta F_u = \frac{\partial F_u}{\partial u} \delta u + \frac{\partial F_u}{\partial \mathbf{u}_1} \cdot \delta \mathbf{u}_1 + \frac{\partial F_u}{\partial \mathbf{u}_2} : \delta \mathbf{u}_2, \quad \delta F_s = \frac{\partial F_s}{\partial s} \delta s.$$

Thus, from the divergence theorem, we have

$$\delta I = \alpha_u \left\langle \delta u, \frac{\partial F_u}{\partial u} - \nabla \cdot \frac{\partial F_u}{\partial \mathbf{u}_1} + \nabla \nabla : \frac{\partial F_u}{\partial \mathbf{u}_2} \right\rangle + \alpha_s \left\langle \delta s, \frac{\partial F_s}{\partial s} \right\rangle + \delta B_n, \tag{19}$$

where

$$\delta B_n = \int_\Gamma \mathbf{n} \cdot \left[\left(\frac{\partial F_u}{\partial \mathbf{u}_1} - \nabla \cdot \frac{\partial F_u}{\partial \mathbf{u}_2} \right) \delta u + \frac{\partial F_u}{\partial \mathbf{u}_2} \cdot \delta \mathbf{u}_1 \right] \mathrm{d}\Gamma. \tag{20}$$

In order to derive (19) and (20), the following vector identity is used:

$$\mathbf{F} : \nabla\nabla u = u \nabla\nabla : \mathbf{F} + \nabla \cdot (\mathbf{F} \cdot \nabla u) - \nabla \cdot ((\nabla \cdot \mathbf{F}) u).$$

Now we consider I_u and I_s, which are related to the equality constraints. Since the differential operators L, M, N are commonly given in the form of the second-order PDEs, we can derive δI_u and δI_s by introducing $\tilde{F}_u = \hat{u} L(u)$ and $\tilde{F}_s = \hat{s}(M(u) + N(s))$ in the above formula. But in the special case when the operators are linear, it is easy to derive the adjoint operator. That is,

$$\delta I_u = \langle \delta \hat{u}, L(u) \rangle + \langle \delta u, L^*(\hat{u}) \rangle + J_L(\delta u, \hat{u}), \tag{21}$$

where L^* is called the adjoint operator and J_L is the conjunct corresponding to the differential operator L, defined by

$$\langle L(u), v \rangle = \langle u, L^*(v) \rangle + J_L(u, v). \tag{22}$$

In the same way, we have

$$\delta I_s = \langle \delta \hat{s}, M(s) + N(u) \rangle + \langle \delta s, M^*(\hat{s}) \rangle + \langle \delta u, N^*(\hat{s}) \rangle + J_M(\delta s, \hat{s}) + J_N(\delta u, \hat{s}). \tag{23}$$

Therefore, from (15), (19), (21), and (23), we have

$$\begin{aligned} \delta I_a = {} & \left\langle \delta u, L^*(\hat{u}) + N^*(\hat{s}) + \alpha_u \left(\frac{\partial F_u}{\partial u} - \nabla \cdot \frac{\partial F_u}{\partial \mathbf{u}_1} + \nabla\nabla : \frac{\partial F_u}{\partial \mathbf{u}_2} \right) \right\rangle \\ & + \left\langle \delta s, M^*(\hat{s}) + \alpha_s \left(\frac{\partial F_s}{\partial s} \right) \right\rangle + \langle \delta \hat{u}, L(u) \rangle + \langle \delta \hat{s}, M(s) + N(u) \rangle \\ & + \delta B_n + J_L(\delta u, \hat{u}) + J_M(\delta s, \hat{s}) + J_N(\delta u, \hat{s}). \end{aligned} \tag{24}$$

From the optimality condition that $\delta I_a = 0$ for arbitrary δu, δs, $\delta \hat{u}$, $\delta \hat{s}$, we obtain Euler–Lagrange equations, given by

$$L(u) = 0, \tag{25}$$

$$M(s) + N(u) = 0, \tag{26}$$

$$L^*(\hat{u}) + N^*(\hat{s}) = \alpha_u \left(- \frac{\partial F_u}{\partial u} + \nabla \cdot \frac{\partial F_u}{\partial \mathbf{u}_1} - \nabla\nabla : \frac{\partial F_u}{\partial \mathbf{u}_2} \right), \tag{27}$$

$$M^*(\hat{s}) = -\alpha_s \left(\frac{\partial F_s}{\partial s} \right). \tag{28}$$

As seen above, the state equations (25) and (26) are recovered by taking the variations to I_a with respect to $\hat{u}$ and $\hat{s}$. The adjoint equations of (27) and (28) are obtained by taking the variations to I_a with respect to u and s. We can see also that the state equations and the adjoint equations are coupled via the derivatives of the integrands in the objective function.

On the other hand, by recalling the fact that the variations of the boundary values are given by $\delta u = 0$ on Γ_e and $\delta u \neq 0$ on Γ_n, we have the natural boundary condition, given by

$$\delta B_n(\mathbf{F}_1, \mathbf{F}_2, \delta u) + J_L(\delta u, \hat{u}) + J_M(\delta s, \hat{s}) + J_N(\delta u, \hat{s}) = 0. \tag{29}$$

The natural boundary condition provides the key equation for u on Γ_n, where u is not specified in the problem. The detailed procedures for obtaining the natural boundary condition become clear in the next section, where each operator is specified.

Generally, the optimization problem has the state constraints in the form of $u \in U$, which can be mathematically handled by projection methods. In most engineering optimization problems, the state constraints can be simplified to inequality constraints. For example, consider

$$u \geq 0, \quad \text{in } \Omega. \tag{30}$$

In this work, we apply the penalty function method and thus the augmented objective function in (13) is modified to

$$I_a = I + I_u + I_s + I_R, \tag{31}$$

where

$$I_R = R \int_\Omega u^2 H(-u)\, \mathrm{d}\Omega. \tag{32}$$

Here, the unit Heaviside step function of $H(-u)$ is defined by

$$H(-u) = \begin{cases} 1, & \text{if } u < 0 \\ 0, & \text{if } u \geq 0. \end{cases} \tag{33}$$

The penalty parameter of R has a large positive value. The limiting case where R has infinitely large value corresponds to the exact consideration of the inequality constraint, but this method is based on the assumption that R with a sufficiently large value is available. By taking the variation of I_R, we have

$$\delta I_R = R \int_\Omega \delta(u^2 H(-u))\, \mathrm{d}\Omega. \tag{34}$$

It is noteworthy that $H(-u)$ is discontinuous at $u = 0$, but that $u^2 H(-u)$ is not only continuous, it is also differentiable due to $\delta(u^2) = 2u\delta u = 0$ at $u = 0$. Thus we have

$$\delta I_R = \langle (2R) u H(-u), \delta u \rangle. \tag{35}$$

Since δI_R is given in the form of domain integration with respect to δu, (27) in the Euler–Lagrange equations is modified to

$$L^*(\hat{u}) + N^*(\hat{s}) = \alpha_u \left(- \frac{\partial F_u}{\partial u} + \nabla \cdot \frac{\partial F_u}{\partial \mathbf{u}_1} - \nabla\nabla : \frac{\partial F_u}{\partial \mathbf{u}_2} \right) - (2R) u H(-u). \tag{36}$$

4. VARIATIONAL FORMULATIONS

4.1. The Optimality System of Equations

The variational method, which is briefly reviewed in Section 3, is now applied to the optimization problem (1) with the equality constraints of (4), (5), (6), and (8). In the same manner as in (13), the augmented objective function of I_a is given by

$$I_a = I + I_{\mathbf{S}} + I_{\mathbf{D}} + I_{\mathbf{v}} + I_u, \tag{37}$$

where

$$I_{\mathbf{S}} = \left\langle \hat{\mathbf{S}}, \mathbf{D} - \frac{1}{2}((\nabla \mathbf{v}) + (\nabla \mathbf{v})^T) \right\rangle, \tag{38}$$

$$I_{\mathbf{D}} = \left\langle \hat{\mathbf{D}}, \mathbf{S} - \frac{1}{(1+\nu)}\mathbf{D} - \frac{\nu(\mathbf{D}:\mathbf{E})}{(1+\nu)(1-2\nu)}\mathbf{E} + \frac{\beta(u-1)}{(1-2\nu)}\mathbf{E} \right\rangle, \tag{39}$$

$$I_{\mathbf{v}} = \langle \hat{\mathbf{v}}, \nabla \cdot \mathbf{S} \rangle, \tag{40}$$

$$I_u = \langle \hat{u}, \nabla^2 u \rangle. \tag{41}$$

Here the definitions of inner products are modified as follows.

$$\langle \mathbf{A}, \mathbf{B} \rangle = \int_\Omega \mathbf{A} : \mathbf{B} \, \mathrm{d}\Omega, \quad \langle \mathbf{u}, \mathbf{v} \rangle = \int_\Omega \mathbf{u} \cdot \mathbf{v} \, \mathrm{d}\Omega, \tag{42}$$

where $\mathbf{A}$ and $\mathbf{B}$ are second-order tensors and $\mathbf{u}$ and $\mathbf{v}$ are vectors. The Lagrange multipliers, $\hat{\mathbf{S}}$, $\hat{\mathbf{D}}$, $\hat{\mathbf{v}}$, and $\hat{u}$, are introduced as the adjoint variables corresponding to the state variables $\mathbf{S}$, $\mathbf{D}$, $\mathbf{v}$, and u, respectively. From the optimality condition $\delta I_a = 0$, we derive the system of equations in the form of coupled PDEs. By applying (19) and (20) to (1) and (2), we have

$$\begin{aligned}
\delta I &= \alpha_u \int_\Omega \delta F_u \, \mathrm{d}\Omega + \alpha_{\mathbf{S}} \int_\Omega \delta F_{\mathbf{S}} \, \mathrm{d}\Omega - \alpha_G \delta G, \\
&= \alpha_u \left\langle \delta u, -\frac{\partial^2 u}{\partial r^2} - \frac{\partial u}{\partial r} \right\rangle + \alpha_{\mathbf{S}} \left\langle \delta \mathbf{S}, \mathbf{S} - \frac{1}{3}(\mathbf{S}:\mathbf{E})\mathbf{E} \right\rangle + \alpha_u \delta B_n - \alpha_G \int_{\Gamma_1} (\mathbf{n} \cdot \nabla \delta u) \, \mathrm{d}\Gamma,
\end{aligned}$$

where

$$\delta B_n = \int_\Gamma \mathbf{n} \cdot \left(2\frac{\partial u}{\partial r}\mathbf{e}_r \right) \delta u \, \mathrm{d}\Gamma. \tag{43}$$

On the other hand, due to the linear properties of the differential operators in the thermal stress problem, it is straightforward to derive each adjoint operator and to rearrange $\delta I_{\mathbf{S}}$, $\delta I_{\mathbf{D}}$, $\delta I_{\mathbf{v}}$, δI_u.

As explained in the previous section, by taking the variations to I_a with respect to adjoint variables of $\hat{\mathbf{S}}$, $\hat{\mathbf{D}}$, $\hat{\mathbf{v}}$, $\hat{u}$, we recover the state equations, given by

$$\mathbf{D} = \frac{1}{2}((\nabla \mathbf{v}) + (\nabla \mathbf{v})^T), \tag{44}$$

$$\mathbf{S} = \frac{1}{(1+\nu)}\mathbf{D} + \frac{\nu}{(1+\nu)(1-2\nu)}(\mathbf{D}:\mathbf{E})\mathbf{E} - \frac{\beta}{(1-2\nu)}(u-1)\mathbf{E}, \tag{45}$$

$$\nabla \cdot \mathbf{S} = 0, \tag{46}$$

$$\nabla^2 u = 0, \tag{47}$$

By taking the variations to I_a with respect to state variables of $\mathbf{S}$, $\mathbf{D}$, $\mathbf{v}$, u, we have the adjoint equations, given by

$$\hat{\mathbf{D}} = \frac{1}{2}((\nabla\hat{\mathbf{v}}) + (\nabla\hat{\mathbf{v}})^T) - \alpha_{\mathbf{S}}\left(\mathbf{S} - \frac{1}{3}(\mathbf{S} : \mathbf{E})\mathbf{E}\right), \tag{48}$$

$$\hat{\mathbf{S}} = \frac{1}{1+\nu}\hat{\mathbf{D}} + \frac{\nu}{(1+\nu)(1-2\nu)}(\hat{\mathbf{D}} : \mathbf{E})\mathbf{E}, \tag{49}$$

$$\nabla \cdot \hat{\mathbf{S}} = 0, \tag{50}$$

$$\nabla^2\hat{u} + \frac{\beta}{(1-2\nu)}(\hat{\mathbf{D}} : \mathbf{E}) = 2\alpha_u\left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r}\right). \tag{51}$$

The inequality constraint of $u \geq 0$ can be imposed when we need to avoid the use of special cooling equipment and reheating of the crystal from a practical point of view. Then, (51) is modified to

$$\nabla^2\hat{u} + \frac{\beta}{(1-2\nu)}(\hat{\mathbf{D}} : \mathbf{E}) = 2\alpha_u\left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r}\right) - (2R)uH(-u), \tag{52}$$

where R is the penalty parameter with a large positive value and H is a unit Heaviside step function defined by (33).

On the other hand, the natural boundary condition is given by

$$\alpha_u \delta B_n - \alpha_G \delta G + J_{\mathbf{v}}(\delta\mathbf{v}, \hat{\mathbf{S}}) + J_{\mathbf{S}}(\delta\mathbf{S}, \hat{\mathbf{v}}) + J_u(\delta u, \hat{u}) = 0, \tag{53}$$

where $J_{\mathbf{v}}$, $J_{\mathbf{S}}$, J_u are the conjuncts corresponding to the differential operators in state equations, defined by

$$J_{\mathbf{v}}(\delta\mathbf{v}, \hat{\mathbf{S}}) = \int_\Gamma \mathbf{n} \cdot (\hat{\mathbf{S}} \cdot \delta\mathbf{v})\, d\Gamma, \tag{54}$$

$$J_{\mathbf{S}}(\delta\mathbf{S}, \hat{\mathbf{v}}) = \int_\Gamma \mathbf{n} \cdot (\delta\mathbf{S} \cdot \hat{\mathbf{v}})\, d\Gamma, \tag{55}$$

$$J_u(\delta u, \hat{u}) = \int_\Gamma \mathbf{n} \cdot (\hat{u}\nabla\delta u - \delta u \nabla\hat{u})\, d\Gamma. \tag{56}$$

By rearranging (53)–(56), we have

$$\begin{aligned} 0 = &\int_\Gamma \mathbf{n} \cdot \left(2\alpha_u \frac{\partial u}{\partial r}\mathbf{e}_r - \nabla\hat{u}\right)\delta u \, d\Gamma + \int_\Gamma \mathbf{n} \cdot (-\alpha_G^* + \hat{u})\nabla(\delta u)\, d\Gamma \\ &+ \int_\Gamma (\mathbf{n} \cdot \hat{\mathbf{S}} \cdot \delta\mathbf{v})\, d\Gamma + \int_\Gamma (\mathbf{n} \cdot \delta\mathbf{S} \cdot \hat{\mathbf{v}})\, d\Gamma, \end{aligned}$$

where

$$\alpha_G^* = \begin{cases} \alpha_G & \text{on } \Gamma_1 \\ 0 & \text{on } \Gamma_2, \Gamma_3, \Gamma_4. \end{cases}$$

Since $\mathbf{n} \cdot \delta\mathbf{S} = 0$ and $\delta\mathbf{v} \neq 0$ along all boundaries, we have

$$\mathbf{n} \cdot \hat{\mathbf{S}} = 0 \quad \text{on } \Gamma_1, \Gamma_2, \Gamma_3. \tag{57}$$

Also from the fact that $\delta u = 0$ on Γ_1, Γ_3 and $\delta u \neq 0$, $\mathbf{n} \cdot \nabla \delta u \neq 0$ on Γ_2, the boundary conditions for $\hat{u}$ are given by

$$\hat{u} = \alpha_G \quad \text{on } \Gamma_1, \tag{58}$$

$$\hat{u} = 0 \quad \text{on } \Gamma_2, \Gamma_3, \tag{59}$$

$$2\alpha_u \frac{\partial u}{\partial r} = \frac{\partial \hat{u}}{\partial r} \quad \text{on } \Gamma_3. \tag{60}$$

As shown in (59), $\hat{u}$ already has the boundary condition of $\hat{u} = 0$ on Γ_3. Thus (60) seems in surplus if it is viewed as a boundary condition for $\hat{u}$. Thus if $\alpha_u \neq 0$, (60) can be taken as a boundary condition for u on Γ_3. But, if $\alpha_u = 0$, we encounter a difficulty. Equation (60) in its form cannot be used as the boundary condition for u even though it should be a key equation for obtaining the boundary condition of u on Γ_3. This point is discussed in Section 5.

4.2. Alternative Formulation of a System of Equations

In this subsection, we discuss a convenient formulation of a system of equations. Although the system of (44)–(51) is natural, it is not convenient in handling. That is because it results in a boundary-value problem on several unknown variables. Thus it is desirable to develop an alternative formulation in which only a minimal number of variables appear in the mathematical boundary-value problem, whereas the remainder are determined afterward by direct calculation. Some detailed discussions are available in the literature [14]. In this work, we consider the displacement formulation. By substituting and eliminating $\mathbf{D}$, $\mathbf{S}$, $\hat{\mathbf{D}}$, and $\hat{\mathbf{S}}$, we have the optimality system of equations in the form of the second-order PDEs, given by

$$\nabla^2 u = 0, \tag{61}$$

$$\nabla^2 \mathbf{v} + \frac{1}{(1-2\nu)} \nabla(\nabla \cdot \mathbf{v}) = 2\beta \frac{(1+\nu)}{(1-2\nu)} \nabla u, \tag{62}$$

$$\nabla^2 \hat{u} + \frac{\beta}{(1-2\nu)} (\nabla \cdot \hat{\mathbf{v}}) = 2\alpha_u \left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} \right) - (2R) u H(-u), \tag{63}$$

$$\nabla^2 \hat{\mathbf{v}} + \frac{1}{(1-2\nu)} \nabla(\nabla \cdot \hat{\mathbf{v}}) = \frac{2\alpha_{\mathbf{S}} \beta}{(1-2\nu)} \nabla u - \frac{2\alpha_{\mathbf{S}}}{3(1-2\nu)} \nabla(\nabla \cdot \mathbf{v}). \tag{64}$$

The stress field and the adjoint stress field can be obtained by the direct calculation of

$$\mathbf{S} = \frac{(\nabla \mathbf{v}) + (\nabla \mathbf{v})^T}{2(1+\nu)} + \frac{\nu(\nabla \cdot \mathbf{v})\mathbf{E}}{(1+\nu)(1-2\nu)} - \frac{\beta(u-1)\mathbf{E}}{(1-2\nu)}, \tag{65}$$

$$\hat{\mathbf{S}} = \frac{(\nabla \hat{\mathbf{v}}) + (\nabla \hat{\mathbf{v}})^T}{2(1+\nu)} + \frac{\nu(\nabla \cdot \hat{\mathbf{v}})\mathbf{E}}{(1+\nu)(1-2\nu)} - \frac{\alpha_{\mathbf{S}} \left(\mathbf{S} - \frac{1}{3}(\mathbf{S} : \mathbf{E})\mathbf{E} \right)}{(1+\nu)}. \tag{66}$$

The conditions of $\mathbf{n} \cdot \mathbf{S} = 0$ and $\mathbf{n} \cdot \hat{\mathbf{S}} = 0$ at Γ_1, Γ_2, Γ_3 are expressed in terms of $\mathbf{v}$ and $\hat{\mathbf{v}}$ as

$$\mathbf{t}\,\mathbf{n}\text{:}\ \frac{((\nabla \mathbf{v}) + (\nabla \mathbf{v})^T)}{2} = 0, \tag{67}$$

$$\mathbf{n}\,\mathbf{n}\text{:}\ \frac{((\nabla \mathbf{v}) + (\nabla \mathbf{v})^T)}{2} + \frac{\nu\,(\nabla \cdot \mathbf{v})}{(1-2\nu)} = \frac{\beta(u-1)}{(1-2\nu)}, \tag{68}$$

$$\mathbf{t}\,\mathbf{n}\text{:}\ \frac{((\nabla \hat{\mathbf{v}}) + (\nabla \hat{\mathbf{v}})^T)}{2} = 0, \tag{69}$$

$$\mathbf{n}\,\mathbf{n}\text{:}\ \frac{((\nabla \hat{\mathbf{v}}) + (\nabla \hat{\mathbf{v}})^T)}{2} + \frac{\nu\,(\nabla \cdot \hat{\mathbf{v}})}{(1-2\nu)} = \frac{\alpha_{\mathbf{S}}(3\beta(u-1) - (\nabla \cdot \mathbf{v}))}{3(1-2\nu)}, \tag{70}$$

where $\mathbf{n}$ and $\mathbf{t}$ denote the unit outward normal and the tangential vectors on the boundaries, respectively. The boundary conditions at the symmetric axis (Γ_4) are given by applying the limiting condition of $r \to 0$ to the governing equations (61)–(64). When the displacement vector and adjoint displacement vector are represented in a cylindrical coordinate system by $\mathbf{v} = v_z \mathbf{e}_z + v_r \mathbf{e}_r$ and $\hat{\mathbf{v}} = \hat{v}_z \mathbf{e}_z + \hat{v}_r \mathbf{e}_r$, we have

$$\Gamma_4 : \frac{\partial u}{\partial r} = \frac{\partial v_z}{\partial r} = v_r = \frac{\partial \hat{u}}{\partial r} = \frac{\partial \hat{v}_z}{\partial r} = \hat{v}_r = 0. \tag{71}$$

On the other hand, the boundary conditions for u and $\hat{u}$ are summarized as

$$\Gamma_1\text{:} \qquad u = 1, \qquad \hat{u} = \alpha_G, \tag{72}$$

$$\Gamma_2\text{:} \qquad u = 0, \qquad \hat{u} = 0, \tag{73}$$

$$\Gamma_3\text{:} \quad 2\alpha_u \frac{\partial u}{\partial r} = \frac{\partial \hat{u}}{\partial r}, \qquad \hat{u} = 0, \tag{74}$$

$$\Gamma_4\text{:} \qquad \frac{\partial u}{\partial r} = 0, \qquad \frac{\partial \hat{u}}{\partial r} = 0. \tag{75}$$

As seen above, we have six coupled PDEs for three state variables u, v_z, v_r and three adjoint variables $\hat{u}$, $\hat{v}_z$, $\hat{v}_r$. It is noteworthy that because the governing equations and the boundary conditions for v_z, v_r, $\hat{v}_z$, $\hat{v}_r$ are given in the differential form, we need additional gauge conditions for the solution uniqueness. In this work, we introduce additional conditions of $v_z = v_r = \hat{v}_z = \hat{v}_r = 0$ at the origin ($z = r = 0$). Physically, the displacement vector $\mathbf{v}$ plus an arbitrary constant means rigid body motion, which does not influence the stress field.

At this point, a comment should be made about problem formulation. Readers may wonder why not use the displacement formulation from the beginning with the state equations (61) and (62). In such a case, the expressions of the objective function and boundary conditions would be very complicated. Consequently a compact treatment such as that shown in Section 4.1 may not be possible, which is why we started with the general expression for the stress distribution, given by (4)–(6), then applied the displacement formulation later.

5. NUMERICAL IMPLEMENTATION

In order to solve the optimality system of equations in (61)–(64) with (67)–(75), we use the finite-difference approximation. When we adopt the operator expressions, the coupling

of unknowns in the differential equations and the boundary conditions may be expressed as

$$
\begin{aligned}
&L^{(1)}\big(x^{(1)}\big) = 0, && B^{(1)}\big(x^{(1)}, x^{(4)}\big) = 0,\\
&L^{(2)}\big(x^{(1)}, x^{(2)}, x^{(3)}\big) = 0, && B^{(2)}\big(x^{(1)}, x^{(2)}, x^{(3)}\big) = 0,\\
&L^{(3)}\big(x^{(1)}, x^{(2)}, x^{(3)}\big) = 0, && B^{(3)}\big(x^{(1)}, x^{(2)}, x^{(3)}\big) = 0,\\
&L^{(4)}\big(x^{(1)}, x^{(4)}, x^{(5)}, x^{(6)}\big) = 0, && B^{(4)}\big(x^{(4)}\big) = 0,\\
&L^{(5)}\big(x^{(1)}, x^{(2)}, x^{(3)}, x^{(5)}, x^{(6)}\big) = 0, && B^{(5)}\big(x^{(1)}, x^{(2)}, x^{(3)}, x^{(5)}, x^{(6)}\big) = 0,\\
&L^{(6)}\big(x^{(1)}, x^{(2)}, x^{(3)}, x^{(5)}, x^{(6)}\big) = 0, && B^{(6)}\big(x^{(1)}, x^{(2)}, x^{(3)}, x^{(5)}, x^{(6)}\big) = 0.
\end{aligned}
$$

Here we set $x^{(1)} = u$, $x^{(2)} = v_z$, $x^{(3)} = v_r$, $x^{(4)} = \hat{u}$, $x^{(5)} = \hat{v}_z$, and $x^{(6)} = \hat{v}_r$ and $L^{(i)}$ and $B^{(i)}$ denote the differential operators for the domain and boundary, respectively. Since the system equations are linear, its discretized version is given in the matrix form of $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$. That is,

$$
\begin{pmatrix}
A^{11} & 0 & 0 & A^{14} & 0 & 0\\
A^{21} & A^{22} & A^{23} & 0 & 0 & 0\\
A^{31} & A^{32} & A^{33} & 0 & 0 & 0\\
A^{41} & 0 & 0 & A^{44} & A^{45} & A^{46}\\
A^{51} & A^{52} & A^{53} & 0 & A^{55} & A^{56}\\
A^{61} & A^{62} & A^{63} & 0 & A^{65} & A^{66}
\end{pmatrix}
\begin{pmatrix} x^{(1)}\\ x^{(2)}\\ x^{(3)}\\ x^{(4)}\\ x^{(5)}\\ x^{(6)} \end{pmatrix}
=
\begin{pmatrix} b^{(1)}\\ b^{(2)}\\ b^{(3)}\\ b^{(4)}\\ b^{(5)}\\ b^{(6)} \end{pmatrix}, \tag{76}
$$

where the coefficient matrix of $\mathbf{A}$ consists of local block matrices A^{ij}. It is noteworthy that the global matrix $\mathbf{A}$ is of the full rank but it is sparse due to trivial off-diagonal matrices. In fact, the existence of nontrivial off-diagonal matrices shows the degree of coupling in the optimality system of equations.

A popular approach to solving $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ in (76) is direct solving methods such as the Gaussian elimination. However, a direct solving method has some fatal weak points. If the system is discretized by n in each direction, then the number of unknowns is n^p, where p is the dimension of the system, i.e., $p = 1, 2, 3$. Since we have three state variables and three adjoint variables, the size of the unknown vector $\mathbf{x}$ is $6n^p$ and thus $\mathbf{A}$ is a $6n^p \times 6n^p$ matrix. For example, in the case of $n = 50$ and $p = 2$, the global matrix $\mathbf{A}$ becomes a matrix about 15000×15000, which may be too huge to be handled by the Gaussian elimination. That is why we want to develop an iterative numerical scheme for solving the optimality system of equations.

In this work, we take advantage of the fact that diagonal block matrices A^{ii} can be inverted with relative ease because the differential equations in (61)–(64) are elliptic with respect to u, $\mathbf{v}$, $\hat{u}$, and $\hat{\mathbf{v}}$, respectively. Based on this idea, we propose an iterative numerical algorithm in the form of a block Jacobi method. First, we rearrange (76) in the form

$$A^{11}x^{(1)} = \tilde{b}^{(1)} = b^{(1)} - A^{14}x^{(4)}, \tag{77}$$

$$A^{22}x^{(2)} = \tilde{b}^{(2)} = b^{(2)} - A^{21}x^{(1)} - A^{23}x^{(3)}, \tag{78}$$

$$A^{33}x^{(3)} = \tilde{b}^{(3)} = b^{(3)} - A^{31}x^{(1)} - A^{32}x^{(2)}, \tag{79}$$

$$A^{44}x^{(4)} = \tilde{b}^{(4)} = b^{(4)} - A^{41}x^{(1)} - A^{45}x^{(5)} - A^{46}x^{(6)}, \tag{80}$$

$$A^{55}x^{(5)} = \tilde{b}^{(5)} = b^{(5)} - A^{51}x^{(1)} - A^{52}x^{(2)} - A^{53}x^{(3)} - A^{56}x^{(6)}, \tag{81}$$

$$A^{66}x^{(6)} = \tilde{b}^{(6)} = b^{(6)} - A^{61}x^{(1)} - A^{62}x^{(2)} - A^{63}x^{(3)} - A^{65}x^{(5)}, \tag{82}$$

which leads to an iterative algorithm to find the unknown variable of $\mathbf{x}$. It is summarized by

Step 1: Guess $x^{(1)}, x^{(2)}, x^{(3)}, x^{(4)}, x^{(5)}, x^{(6)}$.
Step 2: Solve (77)–(82) for $x^{(1)}$–$x^{(6)}$, respectively.
Step 3: Check iterative convergence. If not converged, go back to *Step 2*.

Of course, different algorithms for iterative calculations such as the conjugate gradient method and the successive overrelaxation method can be applied. However, it is noteworthy that the iterative scheme based on the diagonal block matrices has some advantage. Each block diagonal matrix is $n^p \times n^p$. But if we use the alternating directional implicit scheme and the second-order finite-difference approximations, we need to solve only $(n \times n)$ tridiagonal matrices even in three-dimensional problems. Thus, the size of the local coefficient matrices A^{ii} does not cause any difficulties. However, it should be stressed that the above scheme works only when every diagonal block matrix has the full rank to be invertible. Therefore, it is critical to check the condition for each diagonal matrix. This point is discussed again shortly along with the boundary condition.

The existence and uniqueness of the optimal solution is strongly influenced by mathematical properties of quadratic objective function. The objective function in this work consists of three performance measures represented by their weighting parameters α_u, $\alpha_{\mathbf{S}}$, α_G. These three weighting parameters appear in the optimality system of equations (61)–(64), (67)–(75), Among them, α_G appears only at the boundary condition for $\hat{u}$ in (72). Since the boundary condition of $\hat{u} = \alpha_G$ is of Dirichlet type, it can be seen that the weighting parameter α_G appears on the right-hand side of the discretized matrix equation, i.e., $b^{(4)}$ in (76). In fact, when $\alpha_G = 0$, we can see that $\hat{u}$ and $\hat{\mathbf{v}}$ are trivial and u has a linear profile, as is discussed later by investigating the limiting cases of $\alpha_u \to \infty$ or $\alpha_{\mathbf{S}} \to \infty$.

Another weighting parameter $\alpha_{\mathbf{S}}$ appears in (64) and (70). Since (64) is an elliptic-type differential equation with respect to $\hat{\mathbf{v}}$, $\alpha_{\mathbf{S}}$ plays the role of source or sink in the $\hat{\mathbf{v}}$-field. Also $\alpha_{\mathbf{S}}$ in (70) provides a boundary source with respect to $\hat{\mathbf{v}}$. From the viewpoint of the discretized matrix equation (76), $\alpha_{\mathbf{S}}$ appears at off-diagonal block matrices, i.e., A^{51}, A^{52}, A^{53}, A^{61}, A^{62}, A^{63}, which represent the coupling of $\hat{\mathbf{v}}$ with u and $\mathbf{v}$. Thus $\hat{\mathbf{v}}$ becomes trivial in the case of $\alpha_{\mathbf{S}} = 0$. Thereby, the full-rank properties of the diagonal block matrices, A^{55} and A^{66}, are not deteriorated by the weighting parameter $\alpha_{\mathbf{S}}$.

On the other hand, the third weighting parameter α_u appears at (63) and (74). The effects of α_u in (63) are similar to those of $\alpha_{\mathbf{S}}$ in (64). It provides the coupling between u and $\hat{u}$ in the system domain and appears in an off-diagonal block matrix A^{41} of (76). Therefore, we can see that α_u in (63) has no influence on the diagonal block matrices, like $\alpha_{\mathbf{S}}$ and α_G. The only exception arises in (74), i.e.,

$$2\alpha_u \frac{\partial u}{\partial r} = \frac{\partial \hat{u}}{\partial r} \quad \text{at } r = 1. \tag{83}$$

By using the backward difference scheme to second-order accuracy, the finite-difference version of (83) is given by

$$2\alpha_u(3u|_k - 4u|_{k-1} + u|_{k-2}) = (3\hat{u}|_k - 4\hat{u}|_{k-1} + \hat{u}|_{k-2}), \tag{84}$$

where k is the data point index along the radial direction at $r = 1$. In terms of $x^{(1)}$ and $x^{(4)}$, (84) forms the following local matrix equation together with other constraints for u:

$$A^{11}x^{(1)} + A^{14}x^{(4)} = b^{(1)}.$$

It is noteworthy that A^{11} becomes deficient when α_u is trivial. As emphasized previously, it is critical that every diagonal matrix has the full-rank property for the iterative convergence. Thus we need a remedy for the deficient A^{11} when $\alpha_u = 0$. In this work, we derive an auxiliary boundary condition for the crystal surface temperature, which guarantees that A^{11} is of the full rank.

Consider $\alpha_u = 0$, $\alpha_S \neq 0$. For simplicity, we assume the penalty parameter R is equal to zero. In the cylindrical coordinates, the normal and tangential directions at Γ_3 are given by $\mathbf{n} = \mathbf{e}_r$ and $\mathbf{t} = \mathbf{e}_z$, respectively. From the r-component of (70), we have

$$u = 1 + \frac{1}{3\beta}(\nabla \cdot \mathbf{v}) + \frac{(1-2\nu)}{\alpha_S \beta}\left(\frac{\partial \hat{v}_r}{\partial r} + \frac{\nu}{(1-2\nu)}(\nabla \cdot \hat{\mathbf{v}})\right). \tag{85}$$

From (63) with $\hat{u} = \partial \hat{u}/\partial r = 0$ at $r = 1$ (see (74)), we have

$$(\nabla \cdot \hat{\mathbf{v}}) = -\frac{(1-2\nu)}{\beta}\frac{\partial^2 \hat{u}}{\partial r^2}.$$

Thereby, we have

$$u = 1 + \frac{1}{3\beta}(\nabla \cdot \mathbf{v}) + \frac{(1-2\nu)}{\alpha_S \beta}\left(\frac{\partial \hat{v}_r}{\partial r} - \frac{\nu}{\beta}\frac{\partial^2 \hat{u}}{\partial r^2}\right) \quad \text{at } r = 1. \tag{86}$$

Therefore, when $\alpha_u = 0$, (86) can be used as the boundary condition for u on Γ_3 instead of (83) during the iterative computation.

6. RESULTS AND DISCUSSION

6.1. Physical Interpretation of the Weighting Parameters

In this work, the objective function is given by

$$\text{Min. } I = \alpha_u \int_\Omega \left(\frac{\partial u}{\partial r}\right)^2 d\Omega + \alpha_S \int_\Omega \left(\tau_{vM}\right)^2 d\Omega - \alpha_G \int_{\Gamma_1} (\mathbf{n} \cdot \nabla u)\, d\Gamma.$$

As seen above, the objective function consists of three terms: (i) the term for the radial uniformity of the temperature distribution, (ii) the term for the minimization of the von Mises stress distribution, and (iii) the term for the single-crystal growth rate. The weighting parameters of α_u, α_S, α_G represent the relative importance of each term in the objective function. The numerical values of α_u, α_S, α_G should be selected by considering the practical implication.

In order to understand the physical meaning of the weighting parameters,we investigate the limiting cases, where one of three parameters is extremely emphasized. By combining the limiting cases together, we can imagine some general patterns of the optimal solution. For simplicity, the inequality condition of $u \geq 0$ is not considered. First, let us consider the limiting case of $\alpha_u \to \infty$. By taking the limit to the adjoint equation for $\hat{u}$ in (63), we have

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} = 0 \quad \text{in } \Omega,$$

which provides $\partial^2 u/\partial z^2 = 0$ due to $\nabla^2 u = 0$. From $u|_{z=0} = 1$ and $u|_{z=l} = 0$, we obtain the linear solution,

$$u = 1 - z/l. \tag{87}$$

In evaluating the stress field caused by (87), it is more convenient to use the stress formulation rather than the displacement formulation. As shown in Ref. [14], from Hooke's law and the force balance equation, we have the elliptic PDEs with respect to **S**, given by

$$(1+\nu)\nabla^2\mathbf{S} + \nabla\nabla(\mathbf{S}:\mathbf{E}) + \beta\left(\frac{(1+\nu)}{(1-\nu)}\nabla^2 u + \nabla\nabla u\right) = 0.$$

As seen above, only the second and higher order derivatives of the temperature are responsible for the generation of thermal stress. Thereby we have

$$\lim_{\alpha_u\to\infty} u = 1 - \frac{z}{l}, \quad \lim_{\alpha_u\to\infty} \mathbf{S} = 0.$$

It is obvious that another limiting condition of $\alpha_{\mathbf{S}} \to \infty$ results in

$$\lim_{\alpha_{\mathbf{S}}\to\infty} \mathbf{S} = 0.$$

Therefore, we can see that in the case of extremely emphasizing either the radial uniformity or the magnitude of the thermal stress, the optimal state is given by the linear temperature profile and no thermal stress. On the other hand, the other limiting case of $\alpha_G \to \infty$ is different from the former cases. From the fact that the heat transfer rate at the crystal–melt interface, i.e., $-(\partial u/\partial z)_{z=0}$, monotonically increases as α_G increases, we can see that there is no smooth solution when $\alpha_G \to \infty$. Consequently, in a normal situation, the optimal solution is determined by the competition between the linear temperature profile and the infinite axial temperature gradient at the interface.

The order of magnitude of each weighting parameter can be predicted based on the direct calculations of the natural distributions of the variables such as in a simple conduction problem. When the radiative heat transfer at the crystal surface is lumped in the form of Newton's law of cooling with a dimensionless number h (which is the Nusselt number in heat transfer), the heat conduction problem can be handled analytically. Brice [15] derived the series solutions for the heat conduction problem, which are consistent with experimental measurements. He provided some useful single-term approximations to the series solution. When $h \ll 1$ and $z \ll l$, the orders of magnitude of the dimensionless temperature and its derivatives are given by

$$u \sim \frac{(2-hr)}{(2-h)}, \quad \frac{\partial u}{\partial z} \sim -(2h)^{1/2}, \quad \left(\frac{\partial u}{\partial r}\right)^2 \sim \frac{h^2}{(2-h)^2}. \tag{88}$$

By using these approximations, we can roughly predict $O(\alpha_u/\alpha_G)$ when $\alpha_{\mathbf{S}} = 0$. That is,

$$\frac{\alpha_u}{\alpha_G} \sim \frac{-(\partial u/\partial z)}{(\partial u/\partial r)^2}. \tag{89}$$

In the same way, $O(\alpha_{\mathbf{S}}/\alpha_G)$ can be predicted. Völkl [16] presented the relation between the temperature conditions and maximum value of the shear stress, which provides the order of magnitude estimation of the von Mises stress. That is,

$$\tau_{vM} \sim \frac{\beta}{4(1-\nu)}\frac{\partial u}{\partial r}. \tag{90}$$

Thereby we have

$$\frac{\alpha_S}{\alpha_G} \sim \frac{-(\partial u/\partial z)}{(\tau_{vM})^2}. \tag{91}$$

By substituting $\nu = 0.25$, $\beta = 0.01$, $h \ll 1$ into (88)–(91) in the silicon single-crystal growth [4], we have rough estimates of the numerical values of $\alpha_u, \alpha_S, \alpha_G$, summarized by the following.

$$\text{If } \alpha_S = 0, \quad \frac{\alpha_u}{\alpha_G} \gg 10^0. \tag{92}$$

$$\text{If } \alpha_u = 0, \quad \frac{\alpha_S}{\alpha_G} \gg 10^5. \tag{93}$$

When all of three parameters have nonzero values, at least one of the two conditions above should be satisfied in order to obtain physically reasonable solutions.

6.2. Comparison with the Exact Solution for the Simple $\alpha_S = 0$ Case

When the objective function is given without any stress term, i.e., $\alpha_S = 0$ in (1), the optimization problem can be formulated in terms of temperature only. In this case, the exact solution is available. The exact solution can be used for verification of the optimal solution obtained by numerical computations. In addition, the exact solution provides an idea for solving the optimization problem in distributed parameter systems by using parametric search methods, called nonlinear programming methods [17]. As is shown later, the finite truncation of an infinite series of eigenfunctions results in the approximation of infinite function space to finite vector space. Recently, Koh [18] attempted to derive an exact solution in the series form as a part of his thesis, which is mainly related to the optimal control of the point defects distribution in the CZ process. He applied the conjugate gradient method to the discretized optimization problem. Main ideas for an exact solution are the same, but we correct his mistakes in the Bessel function treatments. In this work, we use the exact solution to verify our numerical algorithm, and the implementation results and the strong points of the variational methods as an optimization techniques are briefly discussed.

Consider the optimization problem, given by

$$\text{Min. } I = \alpha_u \int_0^l \int_0^1 \left(\frac{\partial u}{\partial r}\right)^2 r \, dr \, dz + \alpha_G \int_0^1 \left(\frac{\partial u}{\partial z}\bigg|_{z=0}\right) r \, dr,$$

subject to the governing equation of

$$\frac{\partial^2 u}{\partial z^2} + \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} = 0$$

and the boundary conditions of

$$u|_{z=0} = 1, \quad u|_{z=l} = 0, \quad \frac{\partial u}{\partial r}\bigg|_{r=0} = 0, \quad u|_{r=1} = u_s(z),$$

where $u_s(z)$ denotes the unspecified distribution of the surface temperature. By using the complete set of sine functions, we suppose

$$u = 1 - \frac{z}{l} + \sum_{n=1}^{\infty} A_n \frac{I_0(\lambda_n r)}{I_0(\lambda_n)} \sin(\lambda_n z), \tag{94}$$

where I_0 are the modified Bessel functions to zeroth order and the eigenvalues λ_n are given by $\lambda_n = \pi n/l$. Thereby, the unknown function u is represented by an infinite number of the coefficients A_n. The derivatives of u are given by

$$\frac{\partial u}{\partial r} = \sum_{n=1}^{\infty} A_n \lambda_n \frac{I_1(\lambda_n r)}{I_0(\lambda_n)} \sin(\lambda_n z),$$

$$\frac{\partial u}{\partial z} = -\frac{1}{l} + \sum_{n=1}^{\infty} A_n \lambda_n \frac{I_0(\lambda_n r)}{I_0(\lambda_n)} \cos(\lambda_n z).$$

Now that I is a function of A_n, the optimality conditions are given by $\partial I/A_n = 0$ for $n = 1, 2, \ldots \infty$. Thereby, we have

$$A_n = -\frac{\alpha_G}{\alpha_u} \frac{I_0(\lambda_n)}{l\lambda_n} \frac{\int_0^1 r I_0(\lambda_n r)\, dr}{\int_0^1 r I_1^2(\lambda_n r)\, dr}.$$

By using the properties of the modified Bessel function, we obtain

$$A_n = -\left(\frac{\alpha_G}{\alpha_u}\right)\left(\frac{1}{l\lambda_n}\right) \frac{2I_1(\lambda_n) I_0(\lambda_n)}{2I_1(\lambda_n) I_0(\lambda_n) - \lambda_n I_0^2(\lambda_n) + \lambda_n I_1^2(\lambda_n)}. \tag{95}$$

Specific numerical values of A_n are shown in Fig. 2 when the aspect ratio is set by $l = 5$ and the weighting parameters are given by $\alpha_u/\alpha_G = 10$.

As mentioned previously, the representation of the optimal solution in the infinite series form provides an idea for application of parametric search methods. That is, by using the finite truncation of the infinite series, the crystal surface temperature can be approximated

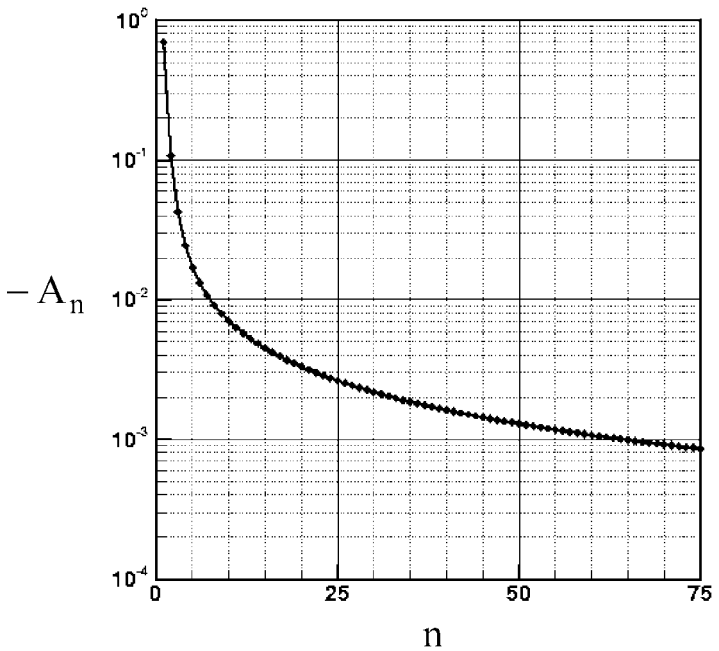

FIG. 2. Optimal set of A_n in the infinite series solution.

by

$$u_s(z) = 1 - \frac{z}{l} + \sum_{n=1}^{N} A_n \sin\left(\frac{n\pi}{l} z\right). \tag{96}$$

Now that the optimization problem is given in terms of A_n, which consists of N-dimensional vector space, various search method techniques are available. For example, the conjugate direction method called Powell's method and the golden section method are recommended for predicting search direction and step size, respectively. In this paper, any further analysis by the search methods is not performed due to space limitation, but the strong and weak points are briefly compared between the search methods and variational methods.

In the search methods, we first solve the state equations by using the initial guess of the control variables, such as A_n in (96). Then we substitute the computed values of state variables into the objective function, which can be evaluated by numerical integration, such as Gaussian quadrature. Now that we have an unconstrained optimization problem as a function of A_n, we can apply the optimization techniques in N-dimensional space. Since the objective function is not given in the quadratic form with respect to A_n, we need iterative computations to obtain the optimal set of A_n. It is noteworthy that in the search method, conventional PDE solution algorithms and optimization techniques are separately applied to solve the state equations and the unconstrained optimization problem, respectively. This advantage should not be taken lightly since many specialized and sophisticated algorithms have been developed for solving them. However, the search method has a critical weak point related to the computational load. In the search method, one iteration claims two different kinds of numerical convergence: One is in solving the state equations in the form of PDEs and the other is in finding the optimal set of A_n, where we need to compute I by numerical integration. Thus the global computation in the search method needs considerable labor. Also, it might be hard to obtain the fully converged solution due to the fluctuations in numerical errors near the solution.

On the other hand, the variational methods adopted in this work have characteristics opposite to the search methods. When the problem is simplified in terms of temperature only, $\hat{\mathbf{v}}$ and $\hat{\mathbf{S}}$ become trivial and $\mathbf{v}$ and $\mathbf{S}$ can be computed explicitly with u; then the optimality system equations are given by

$$\frac{\partial^2 u}{\partial z^2} + \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} = 0,$$

$$\frac{\partial^2 \hat{u}}{\partial z^2} + \frac{\partial^2 \hat{u}}{\partial r^2} + \frac{1}{r}\frac{\partial \hat{u}}{\partial r} = 2\alpha_u \left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r}\right) - 2Ru\,H(-u),$$

where the inequality constraint $u \geq 0$ is added by the penalty term. The boundary conditions are given by

$$u|_{z=0} = 1, \quad \hat{u}_{z=0} = \alpha_G, \quad u|_{z=l} = \hat{u}|_{z=l} = 0,$$

$$\left.\frac{\partial u}{\partial r}\right|_{r=0} = \left.\frac{\partial \hat{u}}{\partial r}\right|_{r=0} = 0, \quad 2\alpha_u \left.\frac{\partial u}{\partial r}\right|_{r=1} = \left.\frac{\partial \hat{u}}{\partial r}\right|_{r=1}, \quad \hat{u}|_{r=1} = 0.$$

As shown above, the problem size doubles by using the adjoint variable $\hat{u}$ and this may be a weak point of the variational approach. However, the numerical evaluation of the objective function is not needed any more because the optimization problem is completely transformed

to differential equations, coupled by means of weighting parameters. Therefore, we need only a PDE solution algorithm, where u and $\hat{u}$ are updated by iterative computation. This point is the most important advantage of the variational approach.

In fact, there are compromised methods between the above two methods. For example, by using the quadratic approximation of the Lagrangian function, the optimization problem can be solved via a series of subproblems which consist of a quadratic objective function and linear constraints [17]. Here each subproblem can be easily handled by quadratic programming techniques with the assumed value of adjoint variable, while the adjoint variable is updated by using the Lagrangian necessary condition of the subproblem. Therefore, the comparison stated above can be regarded as a comparison of two extreme cases of the pure search method and the pure variational method. However, this fact does not significantly change the discussion for the two extreme cases. The pure variational method has some advantage for a certain class of problems in the sense that the numerical evaluation of the objective function is not necessary. Furthermore, the problem of system size increasing by a factor of two is not expected to cause big trouble in view of recent advancements in memory size and computation speed. On the other hand, in the search methods, when exact expressions for the gradient vector and Hessian matrix of the objective function are not available, we have to solve PDEs and we need to perform numerical integrations and numerical differentiations consecutively. These sequential computations may result in the loss of solution accuracy.

In this work, the finite-difference method and the alternative directional iterative method are used to solve the PDEs. The cylindrical crystal ($0 \le z \le l, 0 \le r \le 1$) is discretized by $z = i(l/m)$ and $r = j(1/n)$ with $m = 50$, $n = 20$, and $l = 5$. The penalty function method is used to reflect the technical constraint that the crystal temperature should not be below a certain critical value. In fact, the penalty method for inequality constraints is common in nonlinear programing techniques. However, it is known that the existence of the penalty parameter may result in distortion of the contours of the objective function values. This fact of distortion of contours is the major difficulty with the penalty functions, and thus much effort has been put into developing rules for updating the penalty parameter to increase overall efficiency [17]. On the other hand, in the pure variational approach, the penalty function method does not seem to cause trouble. Numerical errors for u during the iterative computations with different R are shown in Fig. 3a when $\alpha_u = 1.0$, $\alpha_G = 0.1$, and $\alpha_S = 0.0$. The numerical errors are defined as the maximum value of the absolute differences between the values at the present iteration step and those at the previous iteration step. From the figure, we can see that at least there is no bad effect of the penalty parameter R on the solution convergence. Although some fluctuations occur during the iterations, the total number of iterations for the converged solution to the tolerance of 10^{-8} decreases as R increases. Figure 3b shows the crystal surface temperature profiles modified by the penalty parameters. Excessive cooling near the interface for a high axial temperature gradient at $z = 0$ may result in a supercooled region with a temperature below the ambient temperature, i.e., $u < 0$. It is noteworthy that the axial temperature gradient at the crystal–melt interface ($z = 0$) is kept almost constant although the global patterns of the optimal solutions are definitely changed by the penalty parameter R. That is, the temperature distribution near the interface is determined mainly by the ratio of the weighting parameters, α_G/α_u.

It is noteworthy that the penalty function method can be adopted to deal with the constraint to repartition the horizontal section of the crystal, which means that the crystal thermal history, such as crystal cooling rate, is specified in the specific temperature range. In fact, it is

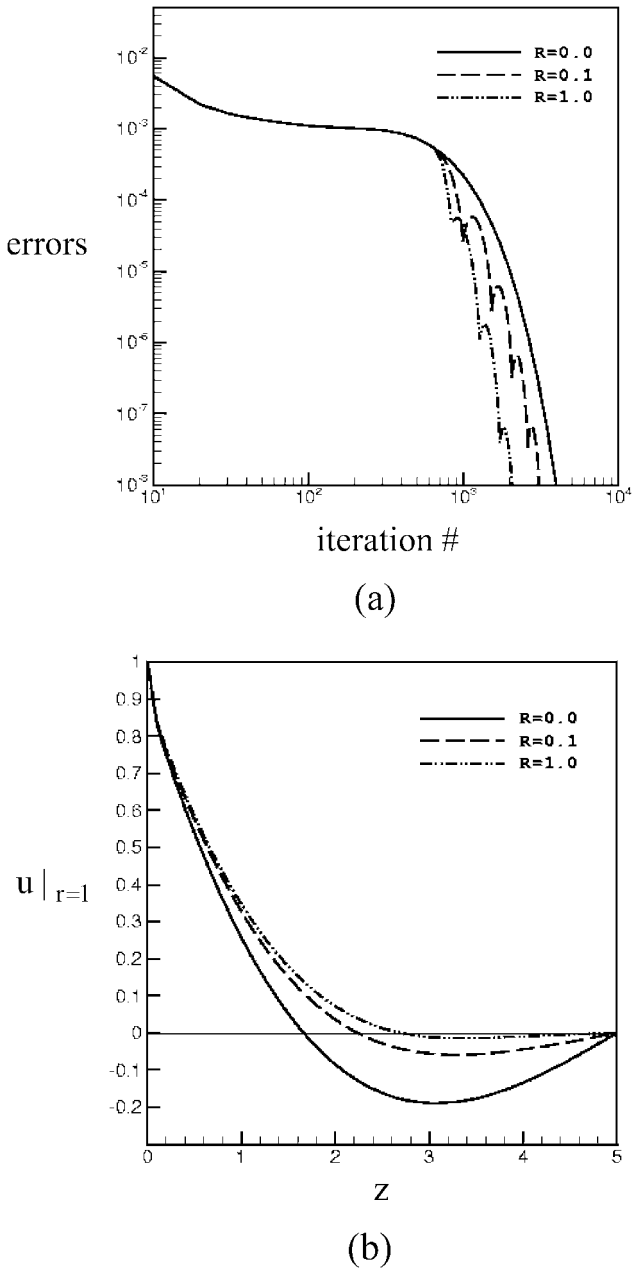

FIG. 3. Results from the variational method with the penalty parameter R when $\alpha_u = 1.0, \alpha_S = 0.0, \alpha_G = 0.1$: (a) numerical errors for u during the iterative computations; (b) the optimal surface temperature profiles.

crucial to understand the role of the thermal history of the growing single-crystal in order to improve the crystal quality. Recently, Takano *et al.* [19] investigated the relationship between grown-in defects and thermal history via direct computation of the temperature field and the measurement of the grown-in defects. They suggested that a specific temperature range exists for the annihilation of the defects. Also they showed that the densities of the defects are strongly influenced by the cooling rate only in the specific temperature region, and their

correlation becomes lower as the cooling rate increases. In our formula, it can be expressed by

$$q_1 \le |\nabla u| \le q_2, \quad \text{for } u_1 \le u \le u_2. \tag{97}$$

For simplicity, consider $|\nabla u| \sim du/dz$, then we have

$$u_1 \le u \le u_2, \quad \text{for } z_1 \le z \le z_2, \tag{98}$$

where $z_1 = z(u_1)$ and $z_2 = z(u_2)$. The inequality condition of (98) can be handled by the penalty function method described by (30)–(36) in Section 3. In this case, (32) is modified by

$$I_R = R \int_{z_1}^{z_2} [(u_1 - u)^2 H(u_1 - u) + (u - u_2)^2 H(u - u_2)]\, \mathrm{d}z. \tag{99}$$

In this paper, any further analysis is not performed due to space limitation. The inequality constraint to repartition the horizontal section of the crystal becomes more important in the point defect problem, where the relative supersaturation of self-interstitials and vacancies should be considered.

In Fig. 4, the effects of R on the contour profiles of the optimal solutions are shown for the case of $\alpha_u = 1.0$, $\alpha_G = 0.1$, and $\alpha_S = 0.0$. From the figure, we can see that the order of magnitude of $O(R) = 10$ is enough to avoid the supercooled region of $u < 0$. From the figure, we can see that the adjoint variable $\hat{u}$ has a singularity at $z = 0, r = 1$ when α_G is not trivial. But the matter does not cause any problem in obtaining the converged solution by iterations.

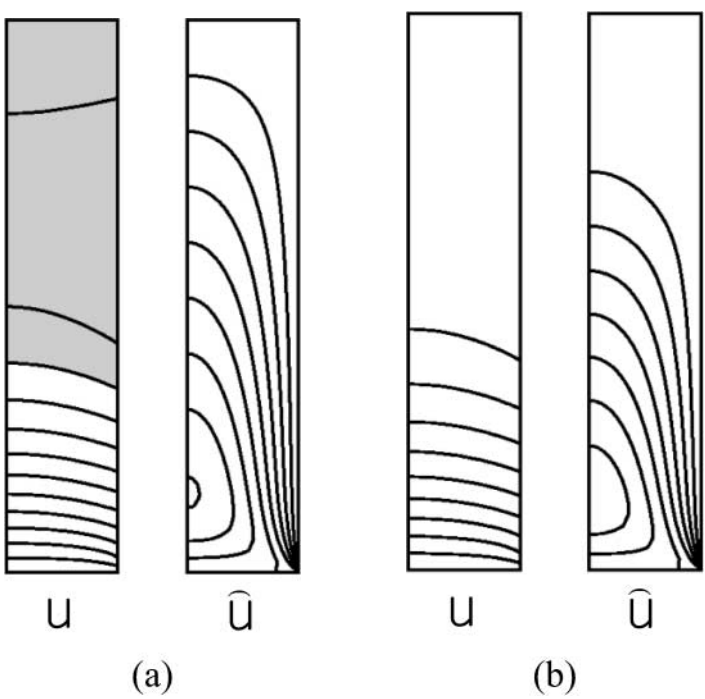

FIG. 4. Effects of the penalty parameter R on the contour profiles of the optimal solutions for the case of $\alpha_u = 1.0, \alpha_S = 0.0, \alpha_G = 0.1$: (a) $R = 0.0$; (b) $R = 70.0$. The difference between the adjacent contours is $\Delta u = 10^{-1}$, $\Delta \hat{u} = 2 \times 10^{-2}$.

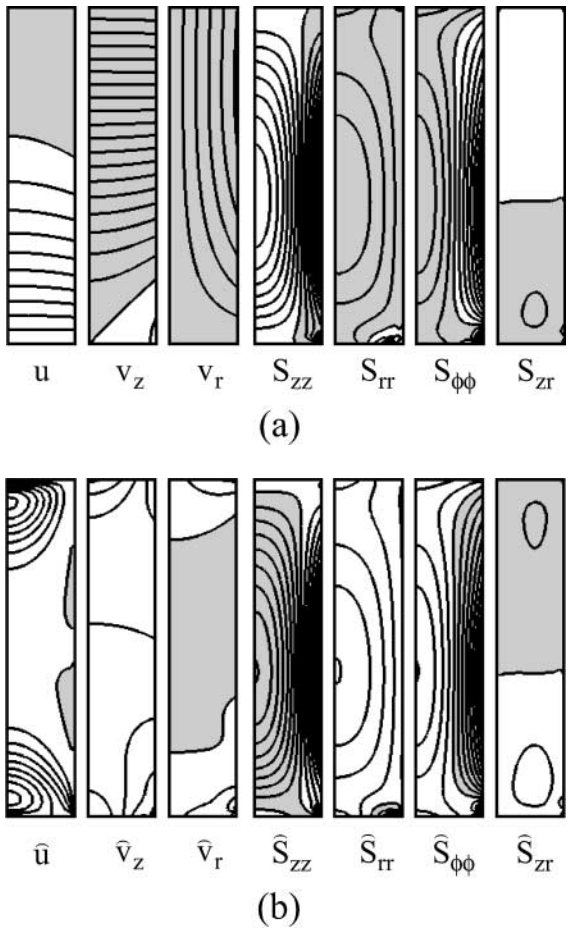

FIG. 5. Prototype results of the optimal solutions when $\alpha_{\mathbf{S}} = 8 \times 10^7$, $\alpha_u = 0.0$, $\alpha_G = 1.0$, $R = 0.0$: (a) state variables; (b) adjoint variables.

6.3. Effects of the Weighting Parameters on the Optimal Solutions

In this subsection, we investigate the effects of the weighting parameters on the optimal solutions. Figure 5 shows a prototype of the results for the optimal solutions when $\alpha_{\mathbf{S}} = 8 \times 10^7, \alpha_u = 0.0, \alpha_G = 1.0, R = 0.0$. In the figures, the differences between the adjacent contours are $\Delta u = \Delta \hat{u} = 10^{-1}$, $\Delta \mathbf{S} = \Delta \hat{\mathbf{S}} = 10^{-5}$, $\Delta \mathbf{v} = \Delta \hat{\mathbf{v}} = 10^{-4}$. In the figure, the variables have negative values in the gray regions. It is interesting that the adjoint variables $\hat{S}_{zz}$, $\hat{S}_{rr}$, $\hat{S}_{\phi\phi}$, and $\hat{S}_{rz}$ have very similar spatial distribution with S_{zz}, S_{rr}, $S_{\phi\phi}$, and S_{rz}, respectively, as if they are mirror images of each other. This fact may result from the similarities in the governing equations and the boundary conditions for $\mathbf{S}$ and $\hat{\mathbf{S}}$.

Figure 6 shows the contours of the optimal distributions of the temperature u and the von Mises stress τ_{vM} for several values of $\alpha_{\mathbf{S}}$ when $\alpha_u = 0.0$, $\alpha_G = 1.0$, $R = 0.0$. That is, the objective function is given by

$$I = \alpha_{\mathbf{S}} \int_{\Omega} (\tau_{vm})^2 \, d\Omega - \int_{\Gamma_1} (\mathbf{n} \cdot \nabla u) \, d\Gamma.$$

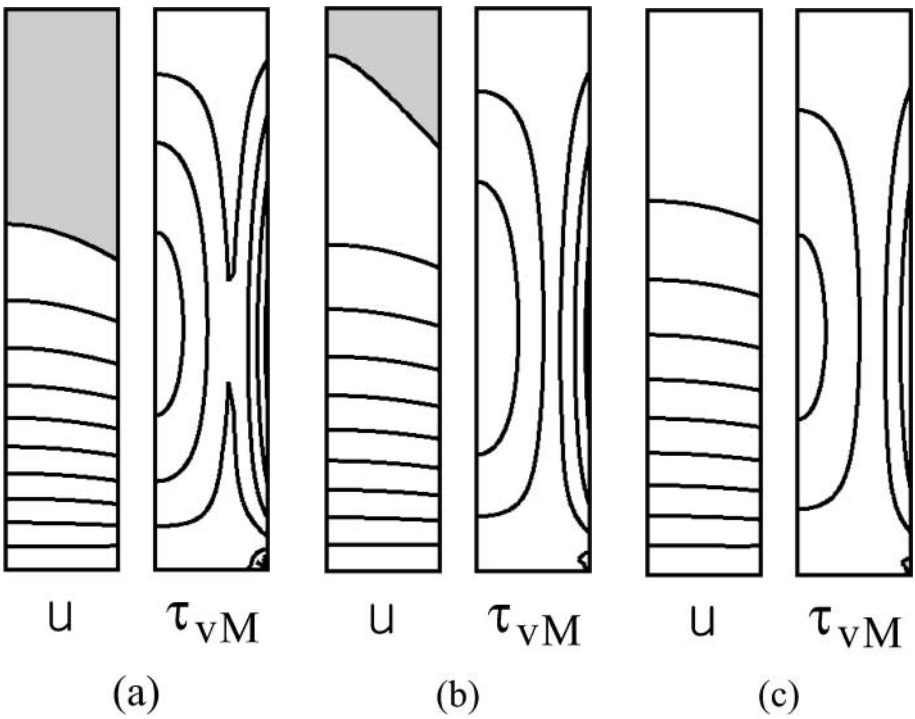

FIG. 6. Effects of α_S on the contour profiles of u and τ_{vM} for the case of $\alpha_u = 0.0$, $\alpha_G = 1.0$, $R = 0.0$: (a) $\alpha_S = 8 \times 10^7$; (b) $\alpha_S = 10 \times 10^7$; (c) $\alpha_S = 12 \times 10^7$.

In the figure, the differences between the adjacent contours Δu and $\Delta \tau_{vM}$ are 10^{-1} and 10^{-5}, respectively. The maximum values for τ_{vM} are 1.7×10^{-4} (Fig. 6a), 1.4×10^{-4} (Fig. 6b), and 1.1×10^{-4} (Fig. 6c) at $(z, r) = (2.5, 1.0)$. From the figure, we can see that the supercooled region ($u < 0$) becomes broader as the weighting parameter α_S decreases. On the other hand, as the weighting parameter α_S increases, the distributions of τ_{vM} become more homogeneous and the maximum value, which occurs at the crystal surface, also decreases. The von Mises stress at a certain point is a qualitative measure for the driving force that induces dislocation, and thus the von Mises stress in the crystal phase should be below a certain critical level, which depends on the material properties of the single crystals. Figure 7 shows the effects of the weighting parameter α_S on the optimal surface temperature profiles when $\alpha_u = 0.0$, $\alpha_G = 0.1$, $R = 0.0$. In the figure, the slope of the axial temperature gradient at the interface ($z = 0$) becomes smaller as the weighting parameter α_S increases. Since the crystal growth rate is directly proportional to the conductive heat transfer at the interface, the crystal growth rate decreases as α_S increases. In other words, in order to reduce the thermal stress below a certain level, it is inevitable to maintain a low crystal growth rate. But the crystal growth rate in each case can be maximized by realizing the crystal surface such as the temperature distributions in Fig. 7.

Figure 8 shows the effects of α_u on the surface temperature profiles when $\alpha_S = 0.0$, $\alpha_G = 1.0$, $R = 0.0$. That is, the objective function is given by

$$I = \alpha_u \int_\Omega \left(\frac{\partial u}{\partial r} \right)^2 d\Omega - \int_{\Gamma_1} (\mathbf{n} \cdot \nabla u)\, d\Gamma.$$

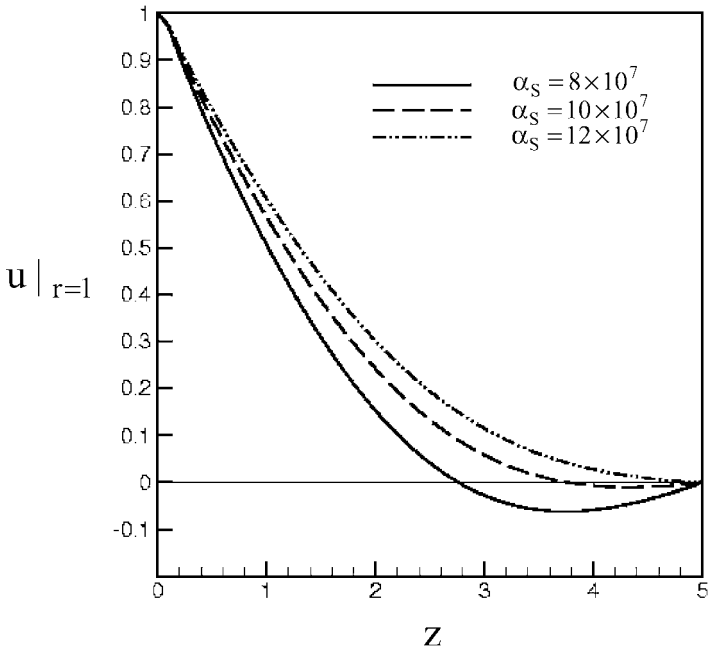

FIG. 7. Effects of α_S on the optimal profiles of the surface temperature when $\alpha_u = 0.0$, $\alpha_G = 1.0$, $R = 0.0$.

It is noteworthy that the effects of α_u are very similar to those of α_S. In fact, this similarity between the effects of α_S and α_u is discussed in Section 5.1 by showing that the optimal solutions are the same, i.e., $u = 1 - z/l$ in the limiting case of either $\alpha_u \to \infty$ or $\alpha_S \to \infty$. Consequently, we can see that the increase in radial uniformity of the temperature

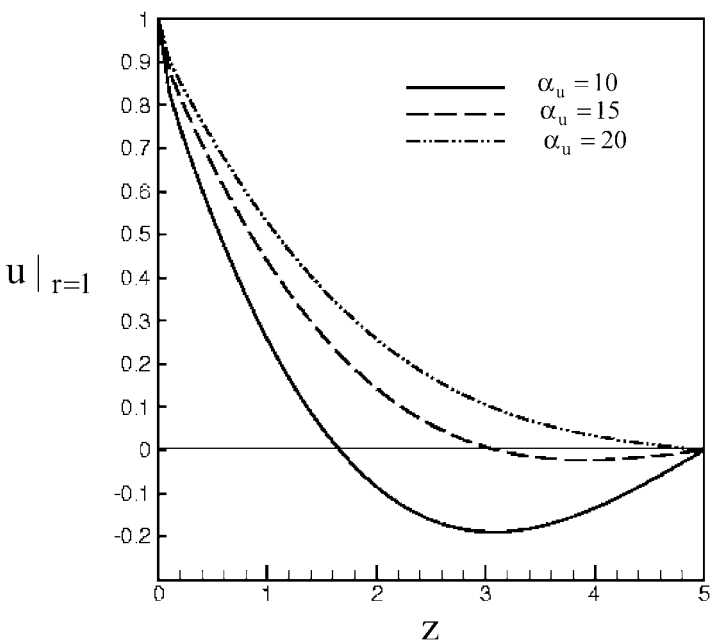

FIG. 8. Effects of α_u on the optimal profiles of the surface temperature when $\alpha_S = 0.0$, $\alpha_G = 1.0$, $R = 0.0$.

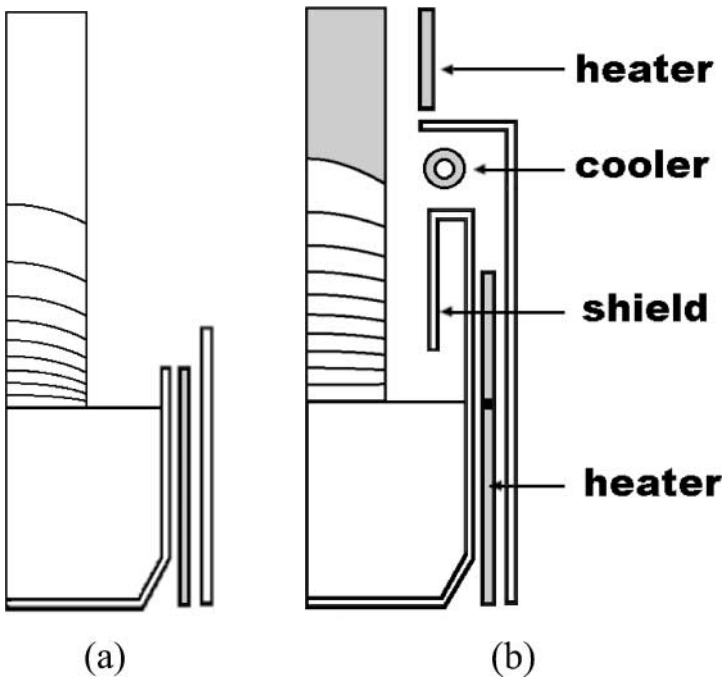

FIG. 9. Schematic for the practical implementation of the optimal solution: (a) a natural growing system without any special equipment for heat transfer control at the crystal surface; (b) the system that can be inferred from the optimal solution.

distribution implies the decrease of the von Mises stress. That is, the radial uniformity of the temperature distribution can be considered an important measure of the thermal stress distribution. This point agrees well with the efforts reported elsewhere. The system configurations and the processing conditions are designed to lower the radial temperature gradients in the crystal phase by using the auxiliary heater, the cold sink, and the thermal shield around the crystal [1, 4].

Figure 9 shows schematic diagram for the practical implementations of the optimal solution to the design of thermal surroundings, such as the thermal shield configuration or heater/cooler positions. Now let us consider a natural growing system without any special equipment for heat transfer control at the crystal surface for comparison with the system that can be inferred from the optimal solution. For simplicity, the radiative heat transfer at the crystal surface in the natural growing system is assumed to be lumped by a linear relationship. That is,

$$-\frac{\partial u}{\partial r} = hu \quad \text{at } r = 1, \tag{100}$$

where h denotes dimensionless heat transfer coefficient. Figure 9a shows the temperature distribution in the natural growing system. The optimal temperature distribution is shown in Fig. 9b for the case of $\alpha_S = 10^8$, $\alpha_G = 1.0$, and $\alpha_u = R = 0.0$.

In fact, it is not possible to prescribe the crystal surface temperature but the surface temperature is determined by the result of heat transfer between the crystal and the ambient surroundings, such as the hot furnace and the cold wall. Among the heat transfer modes between the crystal surface and the heating elements, radiation is essential in the growth

process due to the high melting temperature of most crystalline materials, and thus the thermal shields can be applied to control the radiative heat transfer on the crystal surface. Differently from thermal shields, the design of the auxiliary cooler plays an important role in realizing the supercooled region. Also, an auxiliary heater should be provided together with a cooler to replenish the crystal temperature up to the ambient temperature at the top plate. In this way, the optimal solution obtained in this work may provide us with an insight and with the fundamental information for optimal design by tuning up the thermal shield configuration and the heater/cooler positions, which agree with the results reported elsewhere [1, 3, 4]. It is noteworthy that the work done in this paper can be regarded as part of the global problem, where the position and the power of the heating elements can be a control function. By considering the radiative heat exchange at the crystal surface, another subproblem can be formulated. As the solutions to the subproblem, the position and the power of the heating elements are obtained to provide the crystal surface temperature distribution obtained in this work. In forthcoming works, we will present the results for the optimal thermal surrounding obtained by considering the heat exchanges between the crystal surface and the heating elements.

7. CONCLUSIONS

The optimization of the crystal surface temperature distribution is performed for single-crystal growth in the Czochralski process by considering both the crystalline defects and the single-crystal growth rate. The calculus of variations and the method of Lagrange multiplier is applied and the derived Euler-Lagrange equations are solved by the iterative numerical scheme proposed in this work. Through this work, we have reached the following conclusions.

(i) Based on the calculus of variations, an effective iterative numerical algorithm is proposed for the optimization problem considered in the present work.

(ii) The optimal distributions of the crystal surface temperature obtained from this work may provide an insight into the optimal design of thermal surroundings, such as thermal shield configurations and heater/cooler positions. The optimal solutions agree qualitatively well with previous design efforts reported elsewhere. Thus far, optimal design of processing conditions and the system configurations has mostly been accomplished by the numerical simulations case by case.

(iii) The variational methods in this work can be easily extended to more complicated problems relevant to the Czochralski process, such as the point defect problem.

ACKNOWLEDGMENTS

This work has been partly supported by a grant from the KOSEF through ARC at POSTECH. This work has also been supported by the BK21 program of the Ministry of Education of Korea.

REFERENCES

1. T. Tsukada, K. Kakinoki, M. Hozawa, N. Imaishi, K. Shimamura, and T. Fukuda, Numerical and experimental studies on crack formation in $LiNbO_3$ single crystal, *J. Crystal Growth* **180**, 543 (1997).
2. T. Sinno, R. A. Brown, W. von Ammon, and E. Bornberger, Point defect dynamics and the oxidation-induced stacking-fault ring in Czochralski-grown silicon crystals, *J. Electrochem. Soc.* **145**, 302 (1998).

3. E. Bornberger, and W. von Ammon, The dependence of ring-like distributed stacking faults on the axial temperature gradient of growing Czochralski silicon crystals, *J. Electrochem. Soc.* **143**, 1648 (1996).
4. D. E. Bornside, T. A. Kinney, and R. A. Brown, Minimization of thermoealstic stresses in Czochralski grown silicon: application of the integrated system model, *J. Crystal Growth* **108**, 779 (1991).
5. J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, translated by S. K. Mitter (Springer-Verlag, Berlin, 1971).
6. R. Glowinski and J. L. Lions, Exact and approximate controllability for distributed parameter systems, *Acta Numer.* (1994).
7. J.-W. He, R. Glowinski, R. Metcalfe, A. Nordlander, and J. Periaux, Active control and drag optimization for flow past a circular cylinder: I. Oscillatory cylinder rotation, *J. Comput. Phys.* **163**, 83 (2000), doi:10.1006/jcph.2000.6556.
8. M. Berggren, Numerical solution of a flow-control problem: vorticity reduction by dynamic boundary action, *SIAM J. Sci. Comput.* **19**, 829 (1998).
9. O. Ghattas and J.-H. Bark, Optimal control of two- and three-dimenaional incompressible Navier-Stokes flows, *J. Comput. Phys.* **136**, 231 (1997), doi:10.1006/jcph.1997.5744.
10. L. S. Hou and S. S. Ravindran, Computations of boundary optimal control problems for an electrically conducting fluid, *J. Comput. Phys.* **128**, 319 (1996), doi:10.1006/jcph.1996.0213.
11. A. S. Jordan, R. Caruso, and A. R. Von Neida, A thermoelastic analysis of dislocation generation in pulled GaAs crystals, *Bell System Tech. J.* **59**, 593 (1980).
12. D. Maroudas and R. A. Brown, On the prediction of dislocation formation in semiconductor crystals grown from the melt: analysis of the Haasen model for plastic deformation dynamics, *J. Crystal Growth* **108**, 399 (1991).
13. H. Alexander and P. Haasen, Dislocations and plastic flow in the diamond structure, *Solid State Phys.* **22**, 27 (1968).
14. B. A. Boley and J. H. Weiner, *Theory of Thermal Stresses* (Krieger, Malabar, FL, 1985).
15. J. C. Brice, Analysis of the temperature distribution in pulled crystals, *J. Crystal Growth* **2**, 395 (1968).
16. J. Völkl, Stress in the cooling crystal, in *Handbook of Crystal Growth*, edited by D. T. J. Hurle (North-Holland, Amsterdam, 1994), vol. 2, p. 821.
17. G. V. Reklaitis, A. Ravindran, K. M. Ragsdell, *Engineering Optimization: Methods and Applications* (Wiley, New York, 1983).
18. Y. Koh, Optimization of the surface temperature profile of a single crystal in the Czochralski process, M.S. thesis (POSTECH, South Korea, 1998).
19. K. Takano, M. Iida, E. Iino, M. Kimura, and H. Yamagishi, Relationship between grown-in defects and thermal history during CZ Si crystal growth, *J. Crystal Growth* **180**, 363 (1997).